PROBLEMS IN LOGICAL FORM

Robert C. Moore

SRI International, Menlo Park, CA 94025

I INTRODUCTION

Decomposition of the problem of "language understanding" into manageable subproblems has always posed a major challenge to the development theories of, and systems for, natural-language processing. More or less distinct components are conventionally proposed for handling syntax, semantics, pragmatics, and inference. While disagreement exists as to what phenomena properly belong in each area, and how much or what kinds of interaction there are among these components, there is fairly widespread concurrence as to the overall organization of linguistic processing.

Central to this approach is the idea that the processing of an utterance involves producing an expression or structure that is in some sense a representation of the literal meaning of the utterance. It is often maintained that understanding what an utterance literally means consists in being able to recover this representation. In philosophy and linguistics this sort of representation is usually said to display the _logical_ _form_ of an utterance, so we will refer (somewhat loosely) to the representations themselves as "logical forms."

This paper surveys what we at SRI view as some of the key problems encountered in defining a system of representation for the logical forms of English sentences, and suggests possible approaches to their solution. We will first look at some general issues related to the notion of logical form, and then discuss a number of problems associated with the way information involving certain key concepts is expressed in English. Although our main concern here is with theoretical issues rather than with system performance, this paper is not merely speculative. The DIALOGIC system currently under development in the SRI Artificial Intelligence Center parses English sentences and translates them into logical forms embodying many of the ideas presented here.

II THE NATURE OF LOGICAL FORM

The first question to ask is, why even have a level of logical form? After all, sentences of natural languages are _themselves_ conveyers of meaning; that is what natural languages are for. The reason for having logical forms is to present the literal meanings of sentences more perspicuously than do the sentences themselves. It is sometimes said that natural-language sentences do not "wear their meanings on their sleeves"; logical forms are intended to do exactly that.

From this perspective, the main desideratum for a system of logical form is that its semantics be _compositional_. That is, the meaning of a complex expression should depend only on the meaning of its subexpressions. This is needed for meaning-dependent computational processes to cope with logical forms of arbitrary complexity. If there is to be any hope of maintaining an intellectual grasp of what these processes are doing, they must be decomposable into smaller and smaller meaning-dependent subprocesses operating on smaller and smaller meaningful pieces of a logical form. For instance, if identifying the entities referred to by an utterance is a subprocess of inferring the speaker's intentions, there must be identifiable

pieces of the logical form of the utterance that constitute referring expressions. Having logical forms be semantically compositional is the ultimate expression of this kind of decomposability, as it renders _every_ well-formed subexpression a locus of meaning—and therefore a potential locus of meaning-dependent processing. This is probably a more telling argument for semantic compositionality in designing language-processing systems than in analyzing human language, but it can be reasonably argued that such design principles must be followed by any system, whether natural or artificial, that has to adapt to a complex environment (see [Simon, 1969], especially Chapter 4).[1]

Logical form, therefore, is proposed as a level of representation distinct from surface-syntactic form, because there is apparently no direct way to semantically interpret natural language sentences in a compositional fashion. Some linguists and philosophers have challenged this assumption [Montague, 1974a] [Barwise and Cooper, 1981], but the complexity of their proposed systems and the limited range of syntactic forms they consider leave serious doubt that the logical-form level can be completely bypassed.[2]

Beyond being compositional, it is desirable—though perhaps not essential—that the meaning of a logical form also be independent of the context in which the associated utterance occurs. (The meaning of an expression in natural language, of course, is often context-dependent.) A language-processing system must eventually produce a context-independent representation of what the speaker means by an utterance because the content of the utterance will normally be subjected to further processing after the original context has been lost. In the many cases in which the speaker's intended meaning is simply the literal meaning, a context-independent logical form would give us the representation we need. There is little doubt that some representation of this sort is required. For example, much of our general knowledge of the world is derived from simple assertions of fact in natural language, but our situation would be hopeless if, for every fact we knew, we had to remember the context in which it was obtained before we could use it appropriately. Imagine trying to decide what to do with a tax refund by having to recall whether the topic of conversation was rivers or financial institutions the first time one heard that banks were good places in which to keep money.

As this example suggests, context independence is closely related to the resolution of ambiguity. For any given ambiguity, it is possible to find a case in which the information needed to resolve it is derived from the context of an utterance. Therefore, if the meanings of logical forms are to be context-independent, the system of logical forms must provide distinct, unambiguous representations for all possible readings of an ambiguous utterance. The question remains whether logical form should also provide ambiguous representations to handle cases in which the disambiguating information is obtained later or is simply general world knowledge. The pros and cons of such an approach are far from clear, so we will generally assume only unambiguous logical forms.

Although it is sometimes assumed that a context-independent representation of the literal meaning of a sentence can be derived by using syntactic and semantic knowledge only, some pragmatic factors must also be taken into account. To take a concrete example, suppose the request "Please list the Nobel Prize winners in physics," is followed by the question "Who are the Americans?" The phrase "the Americans" in the second utterance should almost certainly be interpreted as

referring to American winners of the Nobel Prize in physics, rather than all inhabitants or citizens of the United States, as it might be understood in isolation. If the logical form of the utterance is to reflect the intended interpretation, processes that are normally assigned to pragmatics must be used to derive it.

One could attempt to avoid this consequence by representing "the Americans" at the level of logical form as literally meaning all Americans, and have later pragmatic processing restrict the interpretation to American winners of the Nobel Prize in physics. There are other cases, however, for which this sort of move is not available. Consider more carefully the adjective "American." American people could be either inhabitants or citizens of the United States; American cars could be either manufactured or driven in the United States; American food could be food produced or consumed in or prepared in a style indigenous to the United States. In short, the meaning of "American" seems to be no more than "bearing some contextually determined relation to the United States." Thus, there is no definite context-independent meaning for sentences containing modifiers like "American." The same is true for many uses of "have," "of," possessives, locative prepositions [Herskovits, 1980] and compound nominals. The only way to hold fast to the position that the construction of logical-form precedes all pragmatic processing seems to be to put in "dummy" symbols for the unknown relations. This may in fact be very useful in building an actual system,[3] but it is hard to imagine that such a level of representation would bear much theoretical weight.

We will thus assume that a theoretically interesting level of logical form will have resolved contextually dependent definite references, as well as the other "local" pragmatic indeterminacies mentioned. An important consequence of this view is that sentences per se do not have logical forms; only sentences in context do.[4] If we speak loosely of the logical form of a sentence, this is how it should be interpreted.

If we go this far, why not say that all pragmatic processing takes place before the logical form is constructed? That is, why make any distinction at all between what the speaker intends the hearer to infer from an utterance and what the utterance literally means? There are two answers to this. The first is that, while the pragmatic factors we have introduced into the derivation of logical form so far are rather narrowly circumscribed (e.g., resolving definitely determined noun phrases), the inference of speaker intentions is completely open-ended. The problem confronting the hearer is to answer the question, "Why would the speaker say that in this situation?" Practically any relevant knowledge that the speaker and hearer mutually possess [Clark and Marshall, 1981] [Cohen and Perrault, 1981] may be brought to bear in answering this question. From a purely methodological standpoint, then, one would hope to define some more restricted notion of meaning as an intermediate step in developing the broader theory.

Even putting aside this methodological concern, it seems doubtful that a theory of intended meaning can be constructed without a concomitant theory of literal meaning, because the latter notion appears to play an explanatory role in the former theory. Specifically, the literal meaning of an utterance is one of those things from which hearers infer speakers' intentions. For instance, in the appropriate context, "I'm getting cold" could be a request to close a window. The only way for the hearer to understand this as a request, however, is to recover the literal content of the utterance, i.e., that the speaker is getting cold, and to infer from this that the speaker would like him to do something about it.

In summary, the notion of logical form we wish to capture is essentially that of a representation of the "literal meaning in context" of an utterance. To facilitate further processing, it is virtually essential that the meaning of logical-form expressions be compositional and, at the same time, it is highly desirable that they be context-independent. The latter condition requires that a system of logical form furnish distinct representations for the different readings of ambiguous natural-language expressions. It also requires that some limited amount of pragmatic processing be involved in producing those representations. Finally, we note that not all pragmatic factors in the use of language can be reflected in the logical form of an utterance, because some of those factors are dependent on information that the logical form itself provides.

III   FORM AND CONTENT IN KNOWLEDGE REPRESENTATION

Developing a theory of the logical form of English sentences is as much an exercise in knowledge representation as in linguistics, but it differs from most work in artificial intelligence on knowledge representation in one key respect. Knowledge representation schemes are usually intended by their designers to be as general as possible and to avoid commitment to any particular concepts. The essential problem for a theory of logical form, however, is to represent specific concepts that natural languages have special features for expressing information about. Concepts that fall in this category include:

* Events, actions, and processes
* Time and space
* Collective entities and substances
* Propositional attitudes and modalities.

A theory of logical form of natural-language expressions, therefore, is primarily concerned with the content rather than the form of representation. Logic, semantic networks, frames, scripts, and production systems are all different forms of representation. But to say merely that one has adopted one of these forms is to say nothing about content, i.e., what is represented. The representation used in this paper, of course, takes a particular form (higher-order logic with intensional operators) but relatively little will be said about developing or refining that form. Rather, we will be concerned with the question of what particular predicates, functions, operators, and the like are needed to represent the content of English expressions involving concepts in the areas listed above. This project might thus be better described as knowledge encoding to distinguish it from knowledge representation, as it is usually understood in artificial intelligence.

IV   A FRAMEWORK FOR LOGICAL FORM

As mentioned previously, the basic framework we will use to represent the logical form of English sentences is higher-order logic (i.e., higher-order predicate calculus), augmented by intensional operators. At a purely notational level, all well-formed expressions will be in "Cambridge Polish" form, as in the programming language LISP; thus, the logical form of "John likes Mary" will be simply (LIKE JOHN MARY). Despite our firm belief in the principle of semantic compositionality, we will not attempt to give a formal semantics for the logical forms we propose. Hence, our

adherence to that principle is a good-faith intention rather than a demonstrated fact. It should be noted, though, that virtually all the kinds of logical constructs used here are drawn from more formal work of logicians and philosophers in which rigorous semantic treatments are provided.

The only place in which our logical language differs significantly from more familiar systems is in the treatment of quantifiers. Normally the English determiners "every" and "some" are translated as logical quantifiers that bind a single variable in an arbitrary formula. This requires using an appropriate logical connective to combine the contents of the noun phrase governed by the determiner with the contents of the rest of the sentence. Thus "Every P is Q" becomes

(EVERY X (IMPLIES (P X) (Q X))),

and "Some P is Q" becomes

(SOME X (AND (P X) (Q X)))

It seems somewhat inelegant to have to use different connectives to join (P X) and (Q X) in the two cases, but semantically it works.

In an extremely interesting paper, Barwise and Cooper [1981] point out (and, in fact, prove) that there are many determiners in English for which this approach does not work. The transformations employed in standard logic to handle "every" and "some" depend on the fact that any statement about every P or some P is logically equivalent to a statement about everything or something; for example, "Some P is Q" is equivalent to "Something is P and Q." What Barwise and Cooper show is that there is no such transformation for determiners like "most" or "more than half." That is, statements about most P's or more than half the P's cannot be rephrased as statements about most things or more than half of all things.

Barwise and Cooper incorporate this insight into a rather elaborate system modeled after Montague's, so that, among other things, they can assign a denotation to arbitrary noun phrases out of context. Adopting a more conservative modification of standard logical notation, we will simply insist that all quantified formulas have an additional element expressing the restriction of the quantifier. "Most P's are Q" will thus be represented by

(MOST X (P X) (Q X)).

Following this convention gives us a uniform treatment for determined noun phrases:

"Most men are mortal"   (MOST X (MAN X) (MORTAL X))
"Some man is mortal"   (SOME X (MAN X) (MORTAL X))
"Every man is mortal"   (EVERY X (MAN X) (MORTAL X))
"The man is mortal"   (THE X (MAN X) (MORTAL X))
"Three men are mortal"   (3 X (MAN X) (MORTAL X))

Note that we treat "the" as a quantifier, on a par with "some" and "every." "The" is often treated formally as an operator that produces a complex singular term, but this has the disadvantage of not indicating clearly the scope of the expression.

A final point about our basic framework is that most common nouns will be interpreted as relations rather than functions in logical form. That is, even if we know that a person has only one height, we will represent "John's height is 6 feet" as

(HEIGHT JOHN (FEET 6))

rather than

(EQ (HEIGHT JOHN) (FEET 6))[5]

There are two reasons for this: one is the desire for syntactic uniformity; the other is to have a variable available for use in complex predicates. Consider "John's height is more than 5 feet and less than 6 feet." If height is a relation, we can say

(THE L (HEIGHT JOHN L)
     (AND (GT L (FEET 5))
          (LT L (FEET 6)))),

whereas, if length is a function, we would say

(AND (GT (HEIGHT JOHN) (FT 5))
     (LT (HEIGHT JOHN) (FT 6)))

The second variant may look simpler, but it has the disadvantage that (HEIGHT JOHN) appears twice. This is not only syntactically unmotivated, since "John's height" occurs only once in the original English but, what is worse, it may lead to redundant processing later on. Let us suppose that we want to test whether the assertion is true and that determining John's height requires some expensive operation, such as accessing an external database. To avoid doing the computation twice, the evaluation procedure must be much more complex if the second representation is used rather than the first.

V   EVENTS, ACTIONS, AND PROCESSES

The source of many problems in this area is the question of whether the treatment of sentences that describe events ("John is going to New York") should differ in any fundamental way from that of sentences that describe static situations ("John is in New York"). In a very influential paper, Davidson [1967] argues that, while simple predicate/argument notation, such as (LOC JOHN NY), may be adequate for the latter, event sentences require explicit reference to the event as an object. Davidson's proposal would have us represent "John is going to New York" as if it were something like "There is an event which is a going of John to New York":

(SOME E (EVENT E) (GO E JOHN NY))

Davidson's arguments for this analysis are that (1) many adverbial modifiers such as "quickly" are best regarded as predicates of the event, and that (2) it is possible to refer to the event explicitly in subsequent discourse. ("John is going to New York. The trip will take four hours.")

The problem with Davidson's proposal is that for sentences in which these phenomena do not arise, the representation becomes unnecessarily complex. We therefore suggest introducing an event abstraction operator, EVABS, that will allow us to introduce event variables when we need them:

(P X1 ... Xn) <->
     (SOME E (EVENT E) ((EVABS P) E X1 ... Xn))

In simple cases we can use the more straightforward form. The logical form of "John is kissing Mary" would simply be (KISS JOHN MARY). The logical form of "John is gently kissing Mary," however, would be

(SOME E (EVENT E)
     (AND ((EVABS KISS) E JOHN MARY)
          (GENTLE E))))

119

If we let EVABS apply to complex predicates (represented by LAMBDA expressions), we can handle other problems as well. Consider the sentence "Being a parent caused John's nervous breakdown." "Parent" is a relational noun; thus, if John is a parent, he must be the parent of someone, but if John has several children we don't want to be forced into asserting that being the parent of any particular one of them caused the breakdown. If we had PARENT1 as the monadic property of being a parent, however, we could say

(SOME E (EVENT E)
        (AND ((EVABS PARENT1) E JOHN)
             (CAUSE E "John's nervous breakdown")))

We don't need to introduce PARENT1 explicitly, however, if we simply substitute for it the expression,

(LAMBDA X (SOME Y (PERSON Y) (PARENT X Y))),

which would give us

(SOME E (EVENT E)
        (AND ((EVABS (LAMBDA X (SOME Y (PERSON Y)
                                       (PARENT X Y))))
              E
              JOHN)
             (CAUSE E "John's nervous breakdown")))

Another important question is whether actions--that is, events with agents--should be treated differently from events without agents and, if so, should the agent be specially indicated? The point is that, if John kissed Mary, that is something he did, but not necessarily something she did. It is not clear whether this distinction should be represented at the level of logical form or is rather an inference based on world knowledge.

Finally, most AI work on actions and events assumes that they can be decomposed into discrete steps, and that their effects can be defined in terms of a final state. Neither of these assumptions is appropriate for continuous processes; e.g., "The flow of water continued to flood the basement." What the logical form for such statements should look like seems to be a completely open question.[6]

                    VI   TIME AND SPACE

We believe that information about time is best represented primarily by sentential operators, so that the logical form of a sentence like "John is in New York at 2:00" would be something like (AT 2:00 (LOC JOHN NY)). There are two main reasons for following this approach. First, current time can be indicated simply by the lack of any operator; e.g., "John owns Fido" becomes simply (OWNS JOHN FIDO). This is especially advantageous in basically static domains in which time plays a minimal role, so we do not have to put something into the logical form of a sentence that will be systematically ignored by lower-level processing. The other advantage of this approach is that temporal operators can apply to a whole sentence, rather than just to a verb. For instance, in the preferred reading of "The President has lived in the White House since 1800," the referent of "the President" changes with the time contexts involved in evaluating the truth of the sentence. The other reading can be obtained by allowing the quantifier "the" in "the President" to assume a wider scope than that of the temporal operator.

Although we do not strongly distinguish action verbs from stative verbs semantically, there are

syntactic distinctions that must be taken into account before tense can be mapped into time correctly. Stative verbs express present time by means of the simple present tense, while action verbs use the present progressive. Compare:

John kisses Mary (normally habitual)
John is kissing Mary (normally present time)
John owns Fido (normally present time)
John is owning Fido (unacceptable)

This is why (KISS JOHN MARY) represents "John is kissing Mary," rather than "John kisses Mary," which would normally receive a dispositional or habitual interpretation.

What temporal operators will be needed? We will use the operator AT to assert that a certain condition holds at a certain time. PAST and FUTURE will be predicates on points in time. Simple past tense statements with stative verbs, such as "John was in New York," could mean either that John was in New York at some unspecified time in the past or at a contextually specific time in the past:

(SOME T (PAST T) (AT T (LOC JOHN NY)))
(THE T (PAST T) (AT T (LOC JOHN NY)))

(For the second expression to be an "official" logical-form representation, the incomplete definite reference would have to be resolved.) Simple future-tense statements with stative verbs are parallel, with FUTURE replacing PAST. Explicit temporal modifiers are generally treated as additional restrictions on the time referred to. "John was in New York on Tuesday" might be (on at least one interpretation):

(SOME T (AND (PAST T) (DURING T TUESDAY))
        (AT T (LOC JOHN NY))))

For action verbs we get representations of this sort for past and future progressive tenses; e.g., "John was kissing Mary" becomes

(THE T (PAST T) (AT T (KISS JOHN MARY)))

When we use event abstraction to introduce individual events, the interactions with time become somewhat tricky. Since (KISS JOHN MARY) means "John is (presently) kissing Mary," so must

(SOME E (EVENT E) ((EVABS KISS) E JOHN MARY))

Since logically this formal expression means something like "There is (presently) an event which is a kissing of Mary by John," we will interpret the predicate EVENT as being true at a particular time of the events in progress at that time. To tie all this together, "John was kissing Mary gently" would be represented by

(THE T (PAST T)
      (AT T
          (SOME E (EVENT E)
                (AND ((EVABS KISS) E JOHN MARY)
                     (GENTLE E)))))

The major unsolved problem relating to time seems to be reconciling statements that refer to points in time with those that refer to intervals--for instance, "The company earned $5 million in March." This certainly does not mean that at every point in time during March the company earned $5 million. One could invent a representation for sentences about intervals with no particular relation to the representation for sentences about points, but then we would have the difficult task of constantly having to decide which representation is appropriate. This is further complicated by the fact that the same event, e.g. the American Revolution, could be viewed as defining either

120

a point in time or an interval, depending on the time scale being considered.[7] ("At the time of the American Revolution, France was a monarchy," compared with "During the American Revolution, England suffered a decline in trade.") One would hope that there exist systematic relationships between statements about points in time and statements about intervals that can be exploited in developing a logical form for tensed sentences. There is a substantial literature in philosophical logic devoted to "tense logic" [Rescher and Urquhart, 1971] [McCawley, 1981], but almost all of this work seems to be concerned with evaluating the truth of sentences at points, which, as we have seen, cannot be immediately extended to handle sentences about intervals.

We include space under the same heading as time because a major question about space is the extent to which its treatment should parallel that of time. From an objective standpoint, it is often convenient to view physical space and time together as a four-dimensional Euclidean space. Furthermore, there are natural-language constructions that seem best interpreted as asserting that a certain condition holds in a particular place ("In California it is legal to make a right turn on a red light"), just as time expressions often assert that a condition holds at a particular time. The question is how far this analogy between space and time can be pushed.

VII   COLLECTIVE ENTITIES AND SUBSTANCES

Most representation schemes are designed to express information about such discrete, well-individuated objects as people, chairs, or books. Not all objects are so distinct, however; collections and substances seem to pose special difficulties. Collections are often indicated by conjoined noun phrases. If we say "Newell and Simon wrote Human Problem Solving," we do not mean that they each did it individually (cf. "Newell and Simon have PhDs."), rather we mean that they did it as a unit. Furthermore, if we want the treatment of this sentence to be parallel to that of "Quine wrote Word and Object," we need an explicit representation of the unit "Newell and Simon," so that it can play the same role the individual "Quine" plays in the latter sentence. These considerations create difficulties in sentence interpretation because of the possibility of ambiguities between collective and distributed readings. Thus, "Newell and Simon have written many papers," might mean that individually each has written many papers or that they have jointly coauthored many papers. The problems associated with conjoined noun phrases also arise with plural noun phrases and singular noun phrases that are inherently collective. "John, Bill, Joe, and Sam," "the Jones boys," and "the Jones String Quartet" may all refer to the same collective entity, so that an adequate logical-form representation needs to treat them as much alike as possible. These issues are treated in detail by Webber [1978].

The most obvious approach to handling collective entities is to treat them as sets, but standard set theory does not provide quite the right logic. The interpretation of "and" in "the Jones boys and the Smith girls" would be the union of two sets, but in "John and Mary" the interpretation would be constructing a set out of two individuals. Also, the distinction made in set theory between an individual, on one hand, and the singleton set containing the individual, on the other, seems totally artificial in this context. We need a "flatter" kind of structure than is provided by standard set theory. The usual formal treatment of strings is a useful model; there is no distinction made between a character and a string just one character long;

moreover, string concatenation applies equally to strings of one character or more than one. Collective entities have these features in common with strings, but share with sets the properties of being unordered and not having repeated elements.

The set theory we propose has a set formation operator COMB that takes any number of arguments. The arguments of COMB may be individuals or sets of individuals, and the value of COMB is the set that contains all the individual arguments and all the elements of the set arguments; thus,

(COMB A {B C} D {E F G}) = {A B C D E F G}

(The notation using braces is NOT part of the logical-form language; this example is just an attempt to illustrate what COMB means in terms of more conventional concepts.) If A is an individual, (COMB A) is simply A.

We need one other special operator to handle definitely determined plural noun phrases, e.g., "the American ships." The problem is that in context this may refer to some particular set of American ships; hence, we need to recognize it as a definite reference that has to be resolved. Following Weber [1978], We will use the notation (SET X P) to express a predicate on sets that is satisfied by any set, all of whose members satisfy (LAMBDA X P). Then "the P's" would be the contextually determined set, all of whose members are P's:

(THE S ((SET X (P X)) S) ...)

It might seem that, to properly capture the meaning of plurals, we would have to limit the extension of (SET X P) to sets of two or more elements. This is not always appropriate, however. Although "There are ships in the Med," might seem to mean "The set of ships in the Med has at least two members," the question "Are there any ships in the Med?" does not mean "Does the set of ships in the Med have at least two members?" The answer to the former question is yes, even if there is only one ship in the Mediterranean. This suggests that any presupposition the plural carries to the effect that more than one object is involved may be a matter of Gricean implicature ("If he knew there was only one, why didn't he say so?") rather than semantics. Similarly, the plural marking on verbs seems to be just a syntactic reflex, rather than any sort of plural operator. On the latter approach we would have to take "Who killed Cock Robin?" as ambiguous between a singular and plural reading, since singular and plural verb forms would be semantically distinct.

To illustrate the use of our notation, we will represent "Every one of the men who defeated Hannibal was brave." Since no one defeated Hannibal individually, this must be attributed to a collection of men:

```
(SOME T (PAST T)
    (AT T
        (EVERY X (THE S (AND ((SET Y (MAN Y)) S)
                             (DEFEAT S HANNIBAL))
                    (MEMB X S))
            (BRAVE X))))
```

Note that we can replace the plural noun phrase "the men who defeated Hannibal" by the singular collective noun phrase, "the Roman army," as in "Everyone in the Roman army was brave":

```
(SOME T (PAST T)
    (AT T
        (EVERY X (THE S (AND (ARMY S) (ROMAN S))
                    (MEMB X S))
            (BRAVE X))))
```

(AND ((SET Y (MAN Y)) S) (DEFEAT S HANNIBAL))

is replaced by (AND (ARMY S) (ROMAN S)).

Collective entities are not the only objects that are difficult to represent. Artificial intelligence representation schemes have notoriously shied away from mass quantities and substances. ([Hayes, 1978] is a notable exception.) In a sentence like "All Eastern coal contains some sulfur," it seems that "coal" and "sulfur" refer to properties of samples or pieces of "stuff." We might paraphrase this sentence as "All pieces of stuff that are Eastern coal contain some stuff that is sulfur." If we take this approach, then, in interpreting a sentence like "The Universe Ireland is carrying 100,000 barrels of Saudi light crude," we need to indicate that the "piece of stuff" being described is the maximal "piece" of Saudi light crude the ship is carrying.   In other cases, substances seem to be more like abstract individuals, e.g., "Copper is the twenty-ninth element in the periodic table." Nouns that refer to substances can also function as do plural noun phrases in their generic use: "Copper is [antelopes are] abundant in the American southwest."

VIII   PROPOSITIONAL ATTITUDES AND MODALITIES

Propositional   attitudes   and   modalities   are discussed together, because they are both normally treated   as   intensional   sentential   operators.   For instance, to represent "John believes that the Fox is in Naples," we would have an operator BELIEVE that takes "John" as its first argument and the representation of "The Fox is in Naples" as its second argument. Similarly, to represent "The Fox might be in Naples," we could apply an operator POSSIBLE to the representation of "The Fox is in Naples."   This approach works particularly well on a number of problems involving quantifiers. For example, "John believes someone is in the basement" possesses an ambiguity that is revealed by the two paraphrases, "John believes there is someone in the basement" and "There is someone John believes to be in the basement."   As these paraphrases suggest, this distinction is represented by different relative scopes of the belief operator and the existential quantifier introduced by the indefinite pronoun "someone":

(BELIEVE JOHN (SOME X (PERSON X) (LOC X BASEMENT)))

(SOME X (PERSON X) (BELIEVE JOHN (LOC X BASEMENT)))

This approach works very well up to a point, but there are cases it does not handle.   For example, sometimes verbs like "believe" do not take a sentence as an argument, but rather a description of a sentence, e.g., "John believes Goldbach's conjecture." If we were to make "believe" a predicate rather than a sentence operator to handle this type of example, the elegant semantics that has been worked out for "quantifying in" would completely break down. Another alternative is to introduce a predicate TRUE to map a description of a sentence into a sentence that necessarily has the same truth value. Then "John believes Goldbach's conjecture" is treated as if it were "John believes of Goldbach's conjecture that it is true."   This is distinguished in the usual way from "John believes that Goldbach's conjecture (whatever it may be) is true" by reversing the scope of the description "Goldbach's conjecture" and the operator "believe."

The only types of utterances we have tried to represent in logical form to this point are assertions, but of course there are other speech acts as well. The only two we will consider are questions and imperatives (commands).   Since performatives (promises, bets, declarations, etc.) have the same syntactic form as assertions, it appears that they raise no new problems. We will also concern ourselves only with the literal speech act expressed by an utterance. Dealing with indirect speech acts does not seem to change the range of representations needed; sometimes, for example, we may simply need to represent what is literally an assertion as something intended as a command.

For questions, we would like to have a uniform treatment of both the yes/no and WH forms. The simplest approach is to regard the semantic content of a WH question to be a predicate whose extension is being sought.   This does not address the issue of what is a satisfactory answer to a question, but we regard that as part of the theory of speech acts proper, rather than a question of logical form.   We will introduce the operator WHAT for constructing complex set descriptions, which, for the sake of uniformity, we will give the same four-part structure we use for quantifiers. The representation of "What American ships are in the Med?" would roughly be as follows:

(WHAT X (AND (SHIP X) (AMERICAN X))
        (LOC X MED))

WHAT is conveniently mnemonic, since we can represent "who"   as   (WHAT X (PERSON X) ....),   "when"   as (WHAT X (TIME X) ....), and so forth. "How many" questions will be treated as questioning the quantifier. "How many men are mortal?" would be represented as

(WHAT N (NUMBER N)
        (N X (MAN X) (MORTAL X)))

Yes/no questions can be handled as a degenerate case of WH questions by treating a proposition as a 0-ary predicate.   Since the extension of an n-ary predicate is a set of n-tuples, the extension of a proposition would be a set of 0-tuples. There is only one 0-tuple, the empty tuple, so there are only two possible sets of 0-tuples. These are the singleton set containing the empty tuple, and the empty set, which we can identify with the truth values TRUE and FALSE. The logical form of a yes/no question with the proposition P as its semantic content would be (WHAT () TRUE P), or more simply P.

With regard to imperatives, it is less clear what type of semantic object their content should be. We might propose that it is a proposition, but we then have to account for the fact that not all propositions are acceptable as commands. For instance, John cannot be commanded "Bill go to New York." The response that a person can only be "commanded something" he has control over is not adequate, because any proposition can be converted into a command by the verb "make"—e.g., "Make Bill go to New York."

The awkwardness of the phrasing "command someone something" suggests another approach. One commands someone to do something, and the things that are done are actions. If actions are treated as objects, we can define a relation DO that maps an agent and an action into a proposition (See [Moore, 1980]). "John is going to New   York" would   then   be   represented   by (DO JOHN (GO NY)). Actions are now available to be the semantic content of imperatives. The problem with this approach is that we now have to pack into actions all the semantic complexities that can arise in commands—

for instance, adverbial modifiers, which we have treated above as predicates on events ("Go quickly"), quantifiers ("Go to every room in the house"), and negation ("Don't go").

A third approach, which we feel is actually the most promising, is to treat the semantic content of an imperative as being a unary predicate. The force of an imperative is that the person to whom the command is directed is supposed to satisfy the predicate. According to this theory the role of "make" is clear--it converts any proposition into a unary predicate. If the assertion "John is making Bill go to New York" is represented as (MAKE JOHN (GO BILL NY)), we can form a unary predicate by LAMBDA abstraction:

(LAMBDA X (MAKE X (GO BILL NY)),

which would be the semantic content of the command "Make Bill go to New York."

This approach does away with the problem concerning adverbial modifiers or quantifiers in commands; they can simply be part of the proposition from which the predicate is formed. A final piece of evidence favoring this approach over a theory based on the notion of action is that some imperatives have nothing at all to do with actions directly. The semantic content of commands like "Be good" or "Don't be a fool" really does seem to consist exclusively of a predicate.

X CONCLUSION

In a paper that covers such a wide range of disparate topics, it is hard to reach any sweeping general conclusions, but perhaps a few remarks about the nature and current status of the research program are in order. First, it should be clear from the issues discussed that at least as many problems remain in the quest for logical form as have already been resolved. Considering the amount of effort that has been expended upon natural-language semantics, this is somewhat surprising. The reason may be that relatively few researchers have worked in this area for its own sake. Davidson's ideas on action sentences, for instance, raised some very interesting points about logical form-- but the major debate it provoked in the philosophical literature was about the metaphysics of the concept of action, not about the semantics of action sentences. Even when semantics is a major concern, as in the work of Montague, the emphasis is often on showing that relatively well-understood subareas of semantics (e.g., quantification) can be done in a particular way, rather than on attempting to take on really new problems.

An additional difficulty is that so much work has been done in a fragmentary fashion. It is clear that the concept of action is closely related to the concept of time, but it is hard to find any work on either concept that takes the other one seriously. To build a language-processing system or a theory of language processing, however, requires an integrated theory of logical form, not just a set of incompatible fragmentary theories. Our conclusion, then, is that if real progress is to be made on understanding the logical form of natural-language utterances, it must be studied in a unified way and treated as an important research problem in its own right.

NOTES

[1] Although our immediate aim is to construct a theory of natural-language processing rather than truth-conditional semantics, it is worth noting that a system of logical form with a well-defined semantics constitutes a bridge between the two projects. If we have a processing theory that associates English sentences with their logical forms, and if those logical forms have a truth-conditional semantics, then we will have specified the semantics of the English sentences as well.

[2] In other papers (e.g., [Montague, 1974b]), Montague himself uses an intensional logic in exactly the role we propose for logical form--and for much the same reason: "We could ... introduce the semantics of our fragment [of English] directly; but it is probably more perspicuous to proceed indirectly by (1) setting up a certain simple artificial language, that of tensed intensional logic, (2) giving the semantics of that language, and (3) interpreting English indirectly by showing in a rigorous way how to translate it into the artificial language. This is the procedure we shall adopt;..." [Montague, 1974b, p.256].

[3] The DIALOGIC system does build such a representation, or at least components of one, as an intermediate step in deriving the logical form of a sentence.

[4] This suggests that our logical forms are representations of what David Kaplan, in his famous unpublished paper on demonstratives [Kaplan, 1977], calls the content of a sentence, as opposed to its character. Kaplan introduces the content/character distinction to sort out puzzles connected with the use of demonstratives and indexicals. He notes that there are at least two different notions of "the meaning of a sentence" that conflict when indexical expressions are used. If A says to B, "I am hungry," and B says to A, "I am hungry," they have used the same words, but in one sense they mean different things. After all, it may be the case that what A said is true and what B said is false. If A says to B, "I am hungry," and B says to A, "You are hungry," they have used different words, but mean the same thing, that A is hungry. This notion of "meaning different things" or "meaning the same thing" is one kind of meaning, which Kaplan calls "content." There is another sense, though, in which A and B both use the words "I am hungry" with the same meaning, namely, that the same rules apply to determine, in context, what content is expressed. For this notion of meaning, Kaplan uses the term "character." Kaplan's notion, therefore, is that the rules of the language determine the character of a sentence--which, in turn, together with the context of utterance, determines the content. If we broaden the scope of Kaplan's theory to include the local pragmatic indeterminacies we have discussed, it seems that the way they depend on context would also be part of the character of a sentence and that our logical form is thus a representation of the content of the sentence-in-context.

[5] It should be obvious from the example that nouns referring to units of measure--e.g., "feet"--are an exception to the general rule. We treat types of quantities, such as distance, weight, volume, time

duration, etc., as basic conceptual categories. Following Hayes [1979], units such as feet, pounds, gallons, and hours are considered to be functions from numbers to quantities. Thus (FEET 3) and (YARDS 1) denote the same distance. Relations like length, weight, size, and duration hold between an entity and a quantity of an appropriate type. Where a word like "weight" serves in English to refer to both the relation and the quantity, we must be careful to distinguish between them. To see the distinction, note that length, beam, and draft are all relations between a ship and a quantity of the same type, distance. We treat comparatives like "greater than" as multidomain relations, working with any two quantities of the same type (or with pure numbers, for that matter).

6 Hendrix [1973], Rieger [1975], Hayes [1978], and McDermott [1981] have all dealt with continuous processes to some extent, but none of them has considered specifically how language expresses information about processes.

7 This point was impressed upon me by Pat Hayes.

REFERENCES

Barwise, J. and R. Cooper [1981] "Generalized Quantifiers and Natural Language," Linguistics and Philosophy, Vol. 4, No. 2, pp. 159-219 (1981).

Clark, H. and C. Marshall [1981] "Definite Reference and Mutual Knowledge," in Elements of Discourse Understanding: Proceedings of a Workshop on Computational Aspects of Linguistic Structure and Discourse Setting, A. K. Joshi, I. A. Sag, and B. L. Webber, eds. (Cambridge University Press, Cambridge, England, 1981).

Cohen, P. and C. R. Perrault [1981] "Inaccurate Reference," in Elements of Discourse Understanding: Proceedings of a Workshop on Computational Aspects of Linguistic Structure and Discourse Setting, A. K. Joshi, I. A. Sag, and B. L. Webber, eds. (Cambridge University Press, Cambridge, England, 1981).

Davidson, D. [1967] "The Logical Form of Action Sentences," in The Logic of Decision and Action, N. Rescher, ed., pp. 81-95 (University of Pittsburgh Press, Pittsburgh, Pennsylvania, 1967).

Hayes, P. J. [1978] "Naive Physics: Ontology of Liquids," Working Papers, Institute of Semantic and Cognitive Studies, Geneva, Switzerland, (August 1978).

Hayes, P. J. [1979] "The Naive Physics Manifesto," in Expert Systems in the Micro-electronic Age, D. Michie, ed., pp. 242-270 (Edinburgh University Press, Edinburgh, Scotland, 1979).

Hendrix, G. [1973] "Modeling Simultaneous Actions and Continuous Processes," Artificial Intelligence, Vol. 4, Nos. 3, 4, pp. 145-180 (Winter 1973).

Herskovits, A. [1980] "On the Spatial Uses of Prepositions," in Proceedings of the 18th Annual Meeting of the Association for Computational Linguistics, University of Pennsylvania, Philadelphia, Pennsylvania, pp. 1-5 (19-22 June 1980).

Kaplan, D. [1977] "Demonstratives, An Essay on the Semantics, Logic, Metaphysics and Epistemology of Demonstratives and Other Indexicals," unpublished manuscript (March 1977).

McCawley, J. D. [1981] Everything that Linguists Have Always Wanted to Know About Logic but Were Ashamed to Ask (University of Chicago Press, Chicago, Illinois, 1981).

McDermott, D. V. [1981] "A Temporal Logic for Reasoning about Processes and Plans," Research Report 196, Yale University, Department of Computer Science, New Haven, Connecticut (March 1981).

Montague, R. [1974a] "English as a Formal Language," in Formal Philosophy, Selected Papers of Richard Montague, R. H. Thomason, ed., pp. 188-221 (Yale University Press, New Haven, Connecticut, and London, England, 1974).

Montague, R. [1974b] "The Proper Treatment of Quantification in Ordinary English," in Formal Philosophy, Selected Papers of Richard Montague, R. H. Thomason, ed., pp. 188-221 (Yale University Press, New Haven, Connecticut, and London, England, 1974).

Moore, R. C. [1980] "Reasoning About Knowledge and Action," Artificial Intelligence Center Technical Note 191, SRI International, Menlo Park, California (October 1980).

Rescher, N. and A. Urquhart, [1971] Temporal Logic (Springer-Verlag, Vienna, Austria, 1971).

Rieger, C. [1975] "The Commonsense Algorithm as a Basis for Computer Models of Human Memory, Inference, Belief and Contextual Language Comprehension," in Proceedings, Theoretical Issues in Natural Language Processing, Cambridge, Massachusetts, pp. 180-195 (10-13 June 1975).

Simon, H. A. [1969] The Sciences of the Artificial (The MIT Press, Cambridge, Massachusetts, 1969).

Webber, B. L. [1978] "A Formal Approach to Discourse Anaphora," Report No. 3761, Bolt Beranek and Newman, Inc., Cambridge, Massachusetts (May 1978).